RAPIDLY LOCATING SOURCES AND PREDICTING CONTAMINANT DISPERSION IN BUILDINGS

Michael D. Sohn *, Pamela Reynolds, Ashok J. Gadgil and Richard G. Sextro

Indoor Environment Dept., Lawrence Berkeley National Laboratory, Berkeley, CA, USA

ABSTRACT
Contaminant releases in or near a building can lead to significant human exposures unless prompt response measures are taken. However, selecting the proper response depends in part on knowing the source locations, the amounts released, and the dispersion characteristics of the pollutants. We present an approach that estimates this information in real time. It uses Bayesian statistics to interpret measurements from sensors placed in the building yielding best estimates and uncertainties for the release conditions, including the operating state of the building. Because the method is fast, it continuously updates the estimates as measurements stream in from the sensors. We show preliminary results for characterizing a gas release in a three-floor, multi-room building at the Dugway Proving Grounds, Utah, USA.

INDEX TERMS
Multizone modeling, Bayesian statistics, COMIS, Parameter estimation, Sensors

INTRODUCTION
Effective response to unexpected pollutant releases in buildings often requires knowing the source locations, the amounts released, and the duration of the event. However, merely measuring airborne pollutant concentrations using sensors may not reveal this information. Complex airflows typically found in multi-room, multi-floor buildings will often quickly disperse the pollutant throughout the building, leaving insufficient time for the sensors to adequately sample the temporal and spatial profiles of the event. Therefore, sensor measurements must first be interpreted. Moreover, they must be interpreted quickly and continuously as the measurements stream in from the sensors.

Traditional data interpretation algorithms, like optimization, Gibbs sampling, and Kalman filtering, are inadequate for this task. They rely either on simplifying assumptions that are not often met in many indoor pollutant transport systems, or on time-consuming inverse models, which must repeatedly run computationally-intensive fate and transport models after an event has begun (Sohn et al., 2002).

We present an alternative algorithm which uses Bayesian statistics. Our approach succeeds, where traditional methods fail, because it decouples the simulation of predictive fate and transport models from the interpretation of measurements. Time-consuming airflow and pollutant transport predictions and uncertainty estimates are computed prior to a pollutant release. This allows for rapid data interpretation during an event. The technique may be used to estimate the location, magnitude, and duration of the release, to characterize any unknown or variable building or weather conditions, and to predict future pollutant transport in the building. Though Bayesian statistics have been applied to several environmental fields (see

* Contact author email: mdsohn@lbl.gov

Sohn et al. 2002 for a list of recent applications), using it to decouple indoor airflow and transport modeling from data interpretation has, to our knowledge, not been previously reported in the literature.

The objectives of this paper are thus to (1) present our Bayesian approach for interpreting sensor data in real time, and (2) demonstrate the approach by successfully detecting and characterizing a pollutant release in a real multi-floor building.

APPROACH

The Bayesian data interpretation approach is divided into two stages. First, in the *pre-event* or *simulation* stage, all of the time-consuming tasks associated with data interpretation are completed before a pollutant release occurs. A model of the building's indoor airflow and pollutant transport is developed, and input parameters for the model are selected. Any unknown or variable model input, like the location and duration of the pollutant, or the HVAC operating mode, is assigned an uncertainty distribution that describes the probabilistic range of possible values. Generally, wide distributions are assigned given the limited prior information. Next the user generates a library of model realizations by sampling the space of the model parameters using Monte Carlo, or other sampling technique, and predicting airflow and pollutant transport for each set of parameters. Each model realization and model simulation represents a possible building operating condition and pollutant release scenario. Thus, sufficient sampling of the uncertainty distributions is essential to represent the full range of possible building and pollutant release conditions. The resulting library of simulations may consist of several thousand scenarios.

The second stage of the Bayesian approach, *data interpretation*, takes place during a release event. The algorithm compares data streaming in from sensors to the library of pollutant transport predictions using a structured probabilistic method referred to as Bayesian updating (Brand and Small, 1995 and Sohn et al. 2002). Bayes' rule allows us to quickly estimate, and update, the level of agreement between model predictions and the observed data while accounting for the effects of error in the measurements, correlation or averaging of the spatial and temporal data, and any imperfect model representation. See Sohn et al. (2002) for a full description of the technique. To summarize the process, each realization in the library is compared to the data to assess the likelihood that the realization describes the event in progress. A realization with predictions that fit the sensor data well will have a high likelihood estimate. This in turn suggests that the model inputs used to generate that realization in the pre-event simulation stage has high probability of describing the event in progress. Comparing the relative fits for each realization using Bayesian statistics allows us to estimate the best-fitting suite of model inputs and the associated uncertainty.

This second stage of the approach is mathematically simple and can be executed very quickly, much quicker than the rate at which new data will likely arrive from sensors. As long as the original library of simulations provides adequate coverage of the model and input parameter space, the data interpretation during the event can be conducted without further evaluation of the flow and transport models.

APPLICATION

We applied our approach to locate and characterize a pollutant release in a three-floor, multi-room building at the Dugway Proving Grounds, UT. Figure 1 illustrates the building and shows the floorplan of the first floor. The first and second floors each consist of three rooms and a stairwell landing. The third floor consists of a large attic space and a stairwell landing. An air handling unit (AHU) supplies air to the first and second floors and returns air only from the first floor. Researchers from Lawrence Berkeley National Laboratory (LBNL) conducted extensive blower door tests on the building to determine interzonal flow parameters and leakage rates. They also conducted twelve tracer experiments in the building releasing puffs of propylene gas at various release points and under several operating conditions of the AHU. Details of the experimental design were discussed by Sextro et al. (1999).

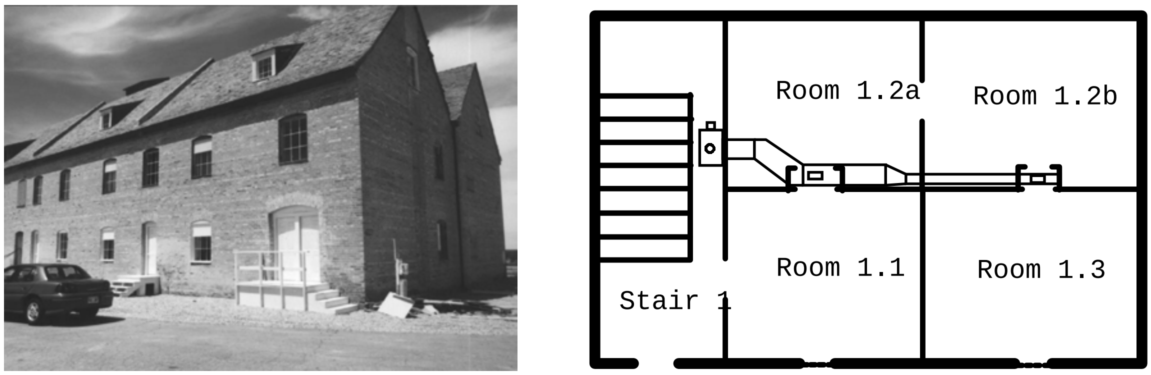

Figure 1: The three-floor building and plan view of the first floor with air handling unit.

Table 1. Uncertainty in the source and building characteristics.

Parameter	*Range*
Source Location	10 locations, consisting of any of the rooms and stairwell. Each location is assumed to be equally likely.
Source Duration	1 sec to 5 min. Log-uniform distribution.
Source Amount	10 to 100 grams. Log-uniform distribution.
Status of Door Positions	3 scenarios, all equally likely: (1) all interior doors open, (2) all interior doors closed, (3) stairwell doors closed, all others open.

In the simulation stage of the Bayesian approach, we hypothesized about the types of pollutant releases that might occur in the building, and assigned uncertainty ranges to the release characteristics (Table 1). We then generated a library of five thousand airflow and pollutant release scenarios by sampling the uncertainty ranges using Latin Hypercube sampling techniques. Airflow and pollutant concentrations in air were predicted for each scenario using the COMIS model, which had been validated previously with experiments conducted in this building (Sextro et al. 1999).

In the data interpretation stage, sensors would be placed at various locations in the building to measure airborne concentrations. The sensors would report the measurements in real-time to the monitoring computer, where they would be interpreted by the algorithm. In the actual experiments conducted at Dugway, propylene detectors were placed in every room and at each stair landing, for a total of eleven sensors. Concentration data were recorded continuously at 20 second intervals. A sample of the data from one of the experiments is shown in Figure 2 at three locations within the building for the first ten minutes of the experiment. In order to test our algorithm, we used data from this experiment, which consisted of a release of ~20 grams of propylene gas into the HVAC return intake in room 1.2a with the HVAC operating. To ensure that our test was valid, we 'blinded' our interpretation algorithm to the actual location and release conditions (see Table 1). The interpretation method was applied for an unknown source location, just as would be the case for a potential release in a real building. For our simulated exercise, we assumed that each data point was reported to the monitoring computer – in this case essentially simultaneously, though the reporting could be asynchronous.

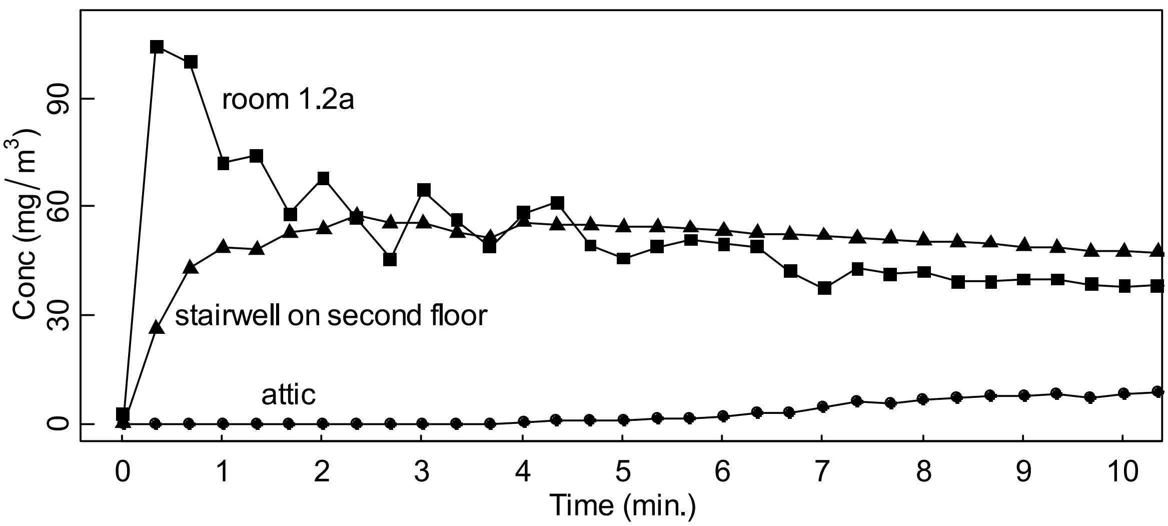

Figure 2: Sensor measurements in room 1.2a (first floor), the stairwell landing on the second floor, and the attic (third floor). Twenty grams of propylene were released in room 1.2a in one second.

Based on the data from all eleven sensors, the interpretation algorithm reports estimates for all of the unknown model inputs, which are updated every 20 seconds. Figures 3a and 3b show the source location probability for two different rooms. It should be noted that at time t=0, the probability of the source location is equal for each room. As can be seen in Figures 3a and 3b, the interpretation algorithm takes less than two minutes to identify the source location (room 1.2a) with high probability.

Since sensors were placed in each room, one might conclude that the location with the highest concentration would always be the release location However, that is correct only when the air is sampled when the source is on and releasing at a constant rate. In this experiment, the release lasted for only one second and the sensors first sampled nineteen seconds after the source stopped. The estimation is further complicated by the AHU quickly dispersing the pollutant throughout the building.

As an example of updating other uncertain model input parameters, Figure 4 shows the algorithm's estimate of the total mass released as it refines that estimate every 20 seconds. The median estimate quickly converges on the correct amount released and uncertainty gradually narrows with each sampling interval. Similar results were found for the other unknowns identified in Table 1, but are not shown.

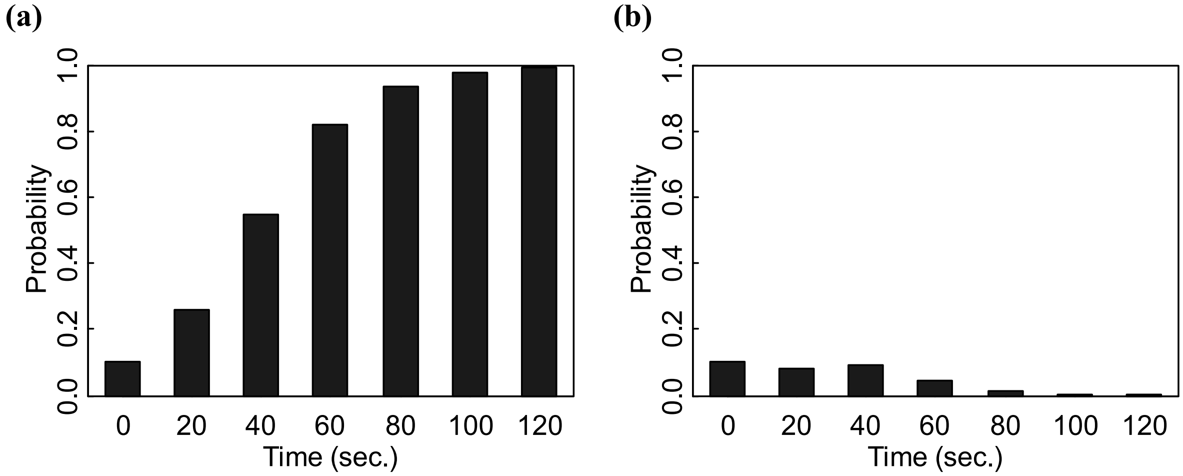

Figure 3: Estimates of the probability that the source is located in room (a) 1.2a and (b) 2.2. At t=0 seconds (before data interpretation begins), the source is assumed be in any of the ten locations with equal probability (i.e., probability = 1/10).

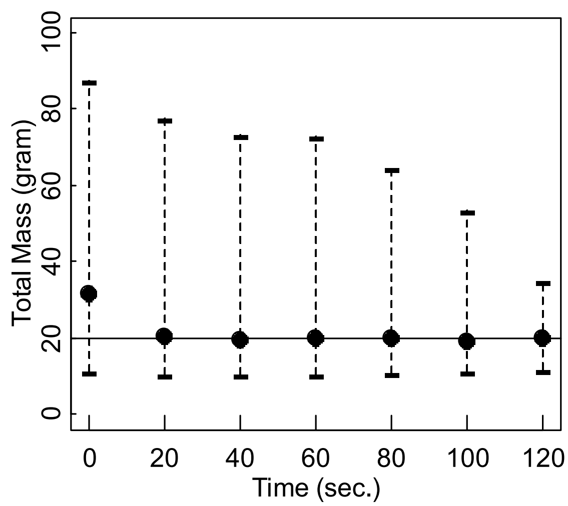

Figure 4: Estimates of the total amount released. At t=0 seconds (before data interpretation begins), the estimate is based on the initial uncertainty defined in Table 1. The solid circle and uncertainty range represent the median and 90 percent confidence interval, respectively, and the horizontal line denotes the actual amount released.

CONCLUSION AND IMPLICATIONS

This paper presents a Bayesian approach for interpreting sensor measurements in real time. It differs from other model parameter estimation methods by decoupling the simulation of airflow and pollutant transport from the interpretation of measurements. This allows us to divide the data interpretation into two parts. The simulation stage completes all of the time consuming tasks, such as development of airflow and pollutant transport models, uncertainty characterization, and simulation of pollutant transport, and compiles the scenario simulations into a library of results. The data interpretation stage quickly accesses this library as data stream in during an event.

We demonstrated the approach by analyzing a gas release in a three-floor building. While the results are preliminary, they illustrate how our approach can quickly interpret data. The approach quickly and correctly identified the source location and the release amount. Though not illustrated here, the algorithm also correctly identified, in less than two minutes, both the duration of the tracer gas release and whether doors were open or closed. Lastly, it correctly predicted the future dispersion of the pollutant in the building.

In future work, we will use our approach to guide sensor deployment. Decoupling data interpretation from model evaluation allows us to compare the performance of many hypothetical sensor operating conditions and sensor locations. Such comparisons could help identify the requirements for a sensor network, including the number, sensitivity, and response time of sensors, based on the desired performance of a data interpretation algorithm in any given building.

ACKNOWLEDGEMENTS
This work was supported by the Office of Non-proliferation and National Security, Chemical and Biological Non-proliferation Program of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. P. Reynolds was supported in part by the U.S.DOE Energy Research Undergraduate Laboratory Fellowship Program.

REFERENCES
Brand, K.P., Small, M.J. 1995. Updating uncertainty in an integrated risk assessment: Conceptual framework and methods. *Risk Analysis* Vol. 15(6) 403-412.

Sextro, R.G., Daisey, J.M., Feustel, H.E., Dickerhoff, D.J., Jump, J. 1999. Comparison of modeled and measured tracer gas concentrations in a multizone building. *Proceedings of the 8th International Conference on Indoor Air Quality and Climate – Indoor Air '99,* Vol 1, pp 696-701. Edinburgh: Indoor Air '99

Sohn, M.D., Reynolds, P, Sing, N., Gadgil, A.J. 2002. Rapidly locating and characterizing pollutant releases in buildings: An application of Bayesian data analysis. LBNL Report 47588. Submitted to *Journal of the Air and Waste Management Association*.